

OPEN

Integrated species–phenon trees: visualizing infraspecific diversity within lineages

Abdullah Khan Zehady¹, Barry G. Fordham^{2*} & James G. Ogg^{3,1}

The unprecedented detail with which contemporary molecular phylogenetics are visualizing infraspecific relationships within living species and species complexes cannot as yet be reliably extended into deep time. Yet paleontological systematics has routinely dealt in (mainly) morphotaxa envisaged in various ways to have been components of past species lineages. Bridging these perspectives can only enrich both. We present a visualization tool that digitally depicts infraspecific diversity within species through deep time. Our integrated species–phenon tree merges ancestor–descendant trees for fossil morphotaxa (phena) into reconstructed phylogenies of lineages (species) by expanding the latter into “species boxes” and placing the phenon trees inside. A key programming strategy to overcome the lack of a simple overall parent–child hierarchy in the integrated tree has been the progressive population of a species–phenon relationship map which then provides the graphical footprint for the overarching species boxes. Our initial case has been limited to planktonic foraminifera via Aze & others’ important macroevolutionary dataset. The tool could potentially be appropriated for other organisms, to detail other kinds of infraspecific granularity within lineages, or more generally to visualize two nested but loosely coupled trees.

Recent advances in molecular systematics are not only building a tree of life but are also allowing the evolutionary structure within a species to be investigated more finely than ever. It is now quite commonplace for species to be revealed as a rich array of clades at multiple taxonomic levels, of difficult-to-assign infraspecific groupings, of cryptic or pseudocryptic species, or of variously fragmented haplotype-phylogeographic populations (follow-up statements in later sections will provide references).

Less well known is that in paleontology an analogous morphotaxonomic granularity within species lineages, long expressed as formal or informal infraspecific morphotypes and the like, is now being portrayed phylogenetically and on a large taxonomic scale — at least in one fossil group. This group is the Cenozoic macroperforate planktonic foraminifera: Aze & others (2011)¹ have proposed for it a species-lineage evolutionary tree comprising 210 lineages, each explicitly incorporating (usually multiple) component morphotaxa.

The macroevolutionary properties of Aze & others’ lineage phylogeny have attracted considerable interest (e.g.²). However, this is not so when it comes to the phylogeny’s microevolutionary ramifications, which in fact do include a granularity within their species lineages implied by the morphotaxa. We contend that this microevolutionary aspect of Aze & others’ phylogeny also deserves attention by evolutionary researchers, especially as most of those species lineages which have survived to the present have now been subjected to molecular analyses, placing planktonic foraminifera in quite a unique position to compare in detail infraspecific diversity from both living and fossil perspectives.

One could nominate important reasons why the focus on Aze & others’ phylogeny has been macro- rather than micro-evolutionary. However, we suggest that there has also been a merely practical hindrance to looking into the content within these lineages — Aze & others’ visuals were simply too complicated to employ for this purpose. This is because they provided the morphotaxonomic content for their species lineages indirectly via a separate evolutionary tree for their morphotaxa, and because the internal timing and the topology of both lineage and morphotaxic trees do not match (deliberately so, but more on that later). So, for the reader to appreciate that content, considerable dexterity is required when comparing corresponding parts of each tree: visual dexterity for

¹Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, USA. ²Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia. ³State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu, Sichuan, China. *email: barry.fordham@anu.edu.au

their tree figures, arithmetic dexterity for their spreadsheet listings (which date and decode their lineages with regard to morphotaxa).

Given the special research potential provided by planktonic foraminifera, we seek to remedy the brain teaser presented by the trees of Aze & others by providing a visualization tool which combines both into a single graphic. In fact, as we will demonstrate, not only does the new integrated tree capture all relevant aspects of each component tree, its additional visual aids delineate when and from where morphological changes are implied to arise both within the species lineage and across proposed speciation events. So it encourages a critical appraisal of the dynamics of morphological evolution in relation to speciation (cladogenesis) and lineage continuity (anagenesis).

Below, we briefly introduce the case study of Aze & others (2011) and we explain why their terminology needs to be adjusted for our tool to bridge both paleontological and living research practices in planktonic foraminifera. We then show that the tool is needed in order to merge two trees (of species lineages and of phena) that are decoupled from each other. In § Results the tool is introduced and applied to the case study, followed by a discussion highlighting mutual learnings it can help broker between paleontological and living research perspectives, as well as other potential applications. § Methods deals with the programming behind the tool and its challenges. But firstly, below we seek to briefly contextualise where visualization within species is at.

Visualizing Intraspecific Relationships

Graphically depicting evolutionary relationships within species—and specifically between entities or taxa below the species level — has been on the agenda even since Darwin. Darwin did so with his tree-like diagram in the *Origin of Species*³ (between pp. 116 and 117) where he envisaged lineal continuity between forms, varieties, subspecies, and species⁴. However, in our contemporary evolutionary lexicon of species as metapopulation lineages⁵, one might presume much of intraspecific usage to be outdated, and so visualizing them to be largely superfluous. That is, until you look a little further. Depictions of varieties and subspecies, for instance, are not hard to find within state-of-the-art taxonomic studies drawn from across the organic world — be they about pathogenic bacteria⁶, reed grasses⁷, endangered crayfish⁸, or titi monkeys⁹. One could in fact get the impression that even the phenon has an assured place in current molecular microbial studies¹⁰.

A key incentive in recent times to visualize relationships within species has been the discrimination increasingly afforded by molecular systematics. Earlier, in both established systematic traditions as well as initial molecular-systematic studies, the focus was and continues to be on superspecific relationships: from specialist research within, say, a family to collaborative ventures aimed at the tree of life. For this context, species are inputs to the study (as such, or as representatives of higher taxa) and their validity is a given. However, the subsequent blossoming of molecular-phylogenetic studies has for the first time allowed in-depth examination of evolution from the gene up. A key component of this research centres around and below the species, with much of the focus in case studies now reframed as the “species complex”^{11–19}. And exciting challenges now involve rethinking of once-given species, including their genetic discrimination²⁰ as well as delineation of their putative evolutionary components.

In this contemporary research environment, a variety of approaches to recognizing and visualizing evolutionary entities around and within the species are being tried. Often, gene trees for species complexes or, say, closely related genera are fully subdivided into clades recognized as species — though the choice and extent of clades is usually manual and pragmatic, influenced by prior systematics¹², and the genetic complexity within and among prior species may raise the level needed to recognize reliable groupings to, say, nominal subgenera¹¹. In other cases, for example: molecular support for prior species may prove poor¹³; geographic variants may cut across prior species¹⁵; or a formal–informal hierarchy of genera, infrageneric clades, species, and unlabeled intraspecific groupings may be viable for most prior species, but there may also be a desiderata of cryptic and pseudocryptic species and species with ambiguous genetic distances¹⁶. For the geographic–genetic differentiation of populations within species or species complexes gained from haplotype phylogeography, mitochondrial-DNA trees are often complemented by annotated networks and ordination of haplogroups^{21–24}. Given the many considerations needed to evaluate gene trees, enhanced visualization approaches have been developed, including embedded display of database information, gene sequences, and phylogeographic maps^{25,26}.

Another layer of interpretative and visualization complexity comes from addressing incongruence between gene and species trees arising from incomplete lineage sorting (deep coalescence), horizontal gene transfer, or gene duplication or loss/extinction events^{27–29}. This applies in particular to very recent speciations or, more generally throughout deep time, when successive speciations are rapid relative to effective population sizes³⁰. So this is especially pertinent for visualizing genetic dynamics within the species, particularly when associated with speciation, potential or eventual. Computational inference of species trees from gene trees³¹ can now employ a rapidly growing array of methods³². When applied to multigene studies, the challenge of visualizing many gene trees^{11,17} may be best guided by the image of a “cloudigram”²⁹ rather than simple lineal trees and, apart from networks and ordination¹⁴, visualizations such as rotatable three-dimensional trees³³ and tanglegrams³⁴ become relevant.

Despite the sophistication which molecular phylogenetics can now bring to differentiation within the species, translating that into phylogenies against deep time has almost always been confined to a simple, and so usually simplistic, projection employing fossil or paleogeographic evidence to date and calibrate selected nodes^{35,36}. So the deep-time history of the molecularly detected intraspecific entities remains largely illusory. Studies which do attempt to detail this history — on groups considered to have an instructive fossil record^{37,38} — will usually need to raise their taxonomic focus for deep time to that of the encompassing species or species groups, providing only a qualitative context to their enhanced depiction of living intraspecific diversity.

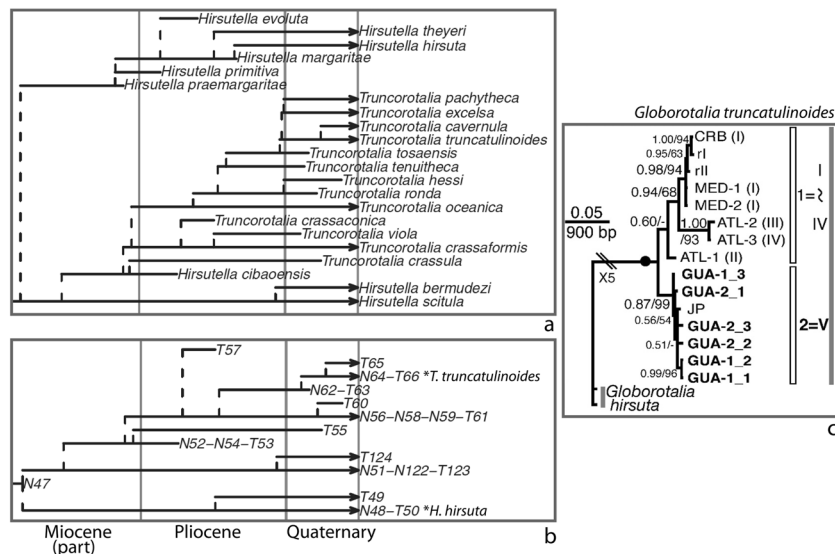


Figure 1. Paleontological and molecular approaches to infraspecific terminology: planktonic foraminifer *Globorotalia/Truncorotalia truncatulinoides*. **(a,b)** A paleontological approach. Portions (after Fig. 5E of¹) of trees of morphospecies [evolutionarily overlapping morphotaxa] **(a)** and lineages [paleobiological species] **(b)**: living Lineage N64-T66 (**T. truncatulinoides* in **b**) includes three living morphospecies (in **a**), *Truncorotalia truncatulinoides*, *T. excelsa*, and *T. pachythea* (*T. cavernula* originated in this lineage, but budded into its own living Lineage T65); for comparison with the outgroup in **(c)**, these portions of the trees embrace living Lineage N48-T50 (**H. hirsuta* in **b**) which includes living morphospecies *Hirsutella hirsuta*; arrows indicate still living. **(c)** A molecular approach. Consensus molecular tree for the morphospecies [cryptic-species complex] *Globorotalia truncatulinoides*, using morphospecies *G. hirsuta* as the outgroup (after Fig. 4b of⁵⁷): two major clades 1 (“1 = I ~ IV”) and 2 (“2 = V”) comprising five genotypes [cryptic species] I – IV and V, respectively. Note: *Globorotalia* versus *Truncorotalia*, *Hirsutella*, etc., are alternative nominal genera, the former preferred for molecular workers’ broader less-settled sweep across living species, the latter for those paleontologists keen to emphasise temporally deep but taxonomically confined lines of descent.

The opportunity now afforded by planktonic foraminifera

A taxonomic area which does offer hope for detailing evolution of molecularly or otherwise delineated entities within species through deep time is that of microfossils³⁹. These phylogenetically diverse groups are united by typically highly rich fossil records from which abundant assemblages can be recovered from tiny samples. Where their fossilisable microscopic parts appear to preserve evolutionary change along stratigraphic sequences, biostratigraphers have typically captured those changes taxonomically, often using a variety of infraspecific labels, formal or informal⁴⁰. So for those microfossil groups that include living representatives, contemporary molecular studies are now providing the potential to compare living infraspecific diversity with that implied through deep time.

A notable example of this opportunity is given by foraminifera, especially the fifty or so living species that are planktonic⁴¹. A growing body of molecular studies on planktonic foraminifera are not merely applying stratigraphic evidence to calibrate their genetic trees against geologic time, but are explicitly exploring a deep-time context for their molecularly detected infraspecific entities^{42–52}. And added to this, their biostratigrapher colleagues have quite recently provided a key macroevolutionary framework for infraspecific diversity through the Cenozoic. For the largest living group of planktonic foraminifera, the macroperforates, Aze & others¹ have proposed phylogenies not just of the usual biostratigraphic taxa but also one which integrates these taxa within whole-species paleobiological lineages, and so depicting deep-time polytypic species directly comparable to living species.

It is Aze & others’ conceptualisation which we consider adds an exciting input into the interchange between living and fossil research, and which forms the case study for our visualization tool. But before introducing the tool, we need to explain why we have avoided their term, morphospecies, and then explain why the tool is needed.

Avoiding “morphospecies”. As already alluded to, Aze & others¹ presented two parallel evolutionary trees for Cenozoic macroperforate planktonic foraminifera: one of biostratigraphic taxa, traditional for micropaleontology, accorded *Genus species* binomina, and termed morphospecies; and another of newly introduced biological-species lineages constructed of mostly multiple morphospecies and given codes. These morphospecies are not chronospecies, they do not subdivide lineages into temporal chunks but rather into segments in morphological space (see p. 195, Text-Fig. 2 in⁵³; pp. 903–905, Fig. 1b in¹; pp. 6–7 in⁵⁴). Despite the pragmatic role that this kind of morphospecies plays in paleontology, in the context of informing and interacting with research into living planktonic foraminifera and its conventions, this usage of “morphospecies” becomes problematic, as seen in the following example.

Aze & others portrayed the later phylogeny of the well-studied living *Truncorotalia* (or *Globorotalia*) *truncatulinoidea*s as two successive Pliocene–Quaternary lineages comprising six morphospecies (Fig. 1a,b; Lineages N62–T63 and N64–T66; morphospecies series beginning with *T. tenuitheca*). As a result, the living *T. truncatulinoidea*s comprised one lineage but three morphospecies, *T. truncatulinoidea*s, *T. excelsa*, and *T. pachythea*. On the other hand, molecular researchers^{43,45,55–59} have consistently considered living *T./G. truncatulinoidea*s a single “morphological species”. This characterisation references an earlier taxonomic tradition which actually included most micropaleontologists whereby only one nominal species, *T./G. truncatulinoidea*s, was recognised for the entire Quaternary lineage (and so also for the present day), whether infraspecific taxa were also delineated^{60–64}, or not^{65–69}.

This disjunct between the terminology of Aze & others and molecular researchers has been further underlined by the latter’s practice of using “morphospecies” as a shorthand for their “morphological species”. So, in current parlance, the living *T./G. truncatulinoidea*s contains three morphospecies in one part of the literature but one morphospecies in another. It is somewhat ironic then that the above-quoted molecular research on *T./G. truncatulinoidea*s has delineated five genotypes, possibly indicative of two genetic lineages, one with four putative species (Fig. 1c). And, though the relationship between these genotypes and the three morphospecies recognised by Aze & others is apparently yet to be addressed, the potential for a coming together of paleontological and molecular concepts for a one-to-one genotype–morphospecies correspondence for these taxa appears poor. Rather, the apparently exclusively subtropical extent (Caribbean, Mediterranean and Canary Islands, west Pacific)^{60,61} of the two ancillary morphospecies, *T. excelsa* and *T. pachythea*, suggests both might be subsumed within only a single genotype which would also include the nominate *T. truncatulinoidea*s (see the distribution of genotype Type II in Fig. 7 of⁵⁶). So, without much more investigation, these three paleontological morphospecies seem unlikely to contribute much to a deep-time context for the five genotypes.

Examples other than *T./G. truncatulinoidea*s would point to similar problems with the “morphospecies” but for different reasons. This is because in molecular studies of living planktonic foraminifera the relationship between the nominal species, usually labeled “morphospecies”, and genetic species is a complicated one^{66,70}; about half of the nominal species analysed each have a one-to-one relationship between nominal and genetic species (though high genetic diversity can nonetheless occur at low genetic levels, as in *Pulleniatina obliquiloculata*⁷⁰); in another half of the nominal species, each contain multiple genotypes suggestive of multiple putative species or, in fact, genetic lineages; the reverse situation of multiple nominal species containing a single genotype is rarer but represented by *Trilobatus*, (Fig. 3 in⁷⁰) apart from the somewhat contrived case of *T./G. truncatulinoidea*s; and finally in some cases a nominal–genetic hierarchy (in either direction) may break down because multiple nominal species are phylogenetically interwoven between multiple genotypes (as in *Globigerinella* and *Globigerinoides ruber* sensu lato⁷⁰).

This discussion suffices to demonstrate the ambiguity and ineffectiveness of “morphospecies” as a term applied to planktonic foraminifera when wanting to integrate current living and paleontological research. Seeking an alternative taxonomic category, the only other option in zoology would be the subspecies, but it would suffer analogous inconsistencies, especially given its important conceptual role in evolutionary practice. So, in applying the case study of Aze & others to our visualization tool, we have retreated to informal practice and replaced Aze & others’ “morphospecies” with Mayr’s use of phena^{71,72}, as applied previously to biostratigraphic taxa in planktonic foraminifer⁶³, more generally^{73,74}, and of course to prokaryotes⁷⁵. This informal paleontological taxonomy then mirrors the emerging informal, though rules-based, system of molecular nomenclature intended to parallel and link to formal nomenclature for planktonic foraminifera and other organisms⁷⁰. And for Aze & others’ other tree, we simply refer to their biological-species lineages as species. Overall, these two adjustments to the terminology employed by Aze & others allow us avoid the very confusing “morphospecies” but otherwise run with the emerging language of molecular research, and comfortably discuss species in an uncontroversial and broadly applicable way.

Why our visualization tool is needed. The above-discussed developments in the study of planktonic foraminifera, from both the living and fossil research communities, can provide a much clearer rationale for extending the living infraspecific diversity of planktonic foraminifera into deep time. There is, however, a practical issue in applying Aze & others’ phylogenies to this purpose. Their separate trees of species (biological-species lineages) and phena (biostratigraphic “morphospecies”) have topologies which do not correspond in a straightforward way. This is because the timing they interpreted for the emergence of species in deep time, though influenced by the stratigraphic ranges of related phena, was based on different evidence — the tracing of morphological clusters of collections of specimens (not taxa, formal or informal) along stratigraphic sequences (p. 195, Text-Fig. 3 in⁵³; Fig. 2 in⁴; § Ancestor–descendant relationships in⁵⁴). This methodology disarticulates the origins of species from those of phena, and so allows phenon ranges to pass unbroken from an ancestral species to one of its descendant species (see morphospecies m3 in Fig. 1b of⁴; note also in Fig. 1a,b herein that, for instance, phenon *T. truncatulinoidea*s appears in the uppermost Pliocene so, when Lineage N64–T66 buds off from its ancestral Lineage N62–T63 in the early Quaternary, that phenon would carry over from its membership of the latter to that of the former). This breakdown in the simple species–phenon hierarchy also underlines the inappropriateness of employing, for example, nominal subspecies for these biostratigraphic taxa.

Making sense of these biostratigraphers’ out-of-synch evolutionary trees of species and phena turns out to be a visual brain-teaser. In this paper our visualization tool provides a graphical solution to this. Our integrated species–phenon tree takes the species tree and widens each species range line into a “species box” within which is displayed that species’ algorithmically determined portion of the tree of phena. Within a species box, symbols attached to the range lines of the phenon tree provide guides to both the origins of the phenon phylogeny from ancestral species and its continuance into descendant species. And all this is accomplished programmatically from a single extra data link between the two trees’ datasets.

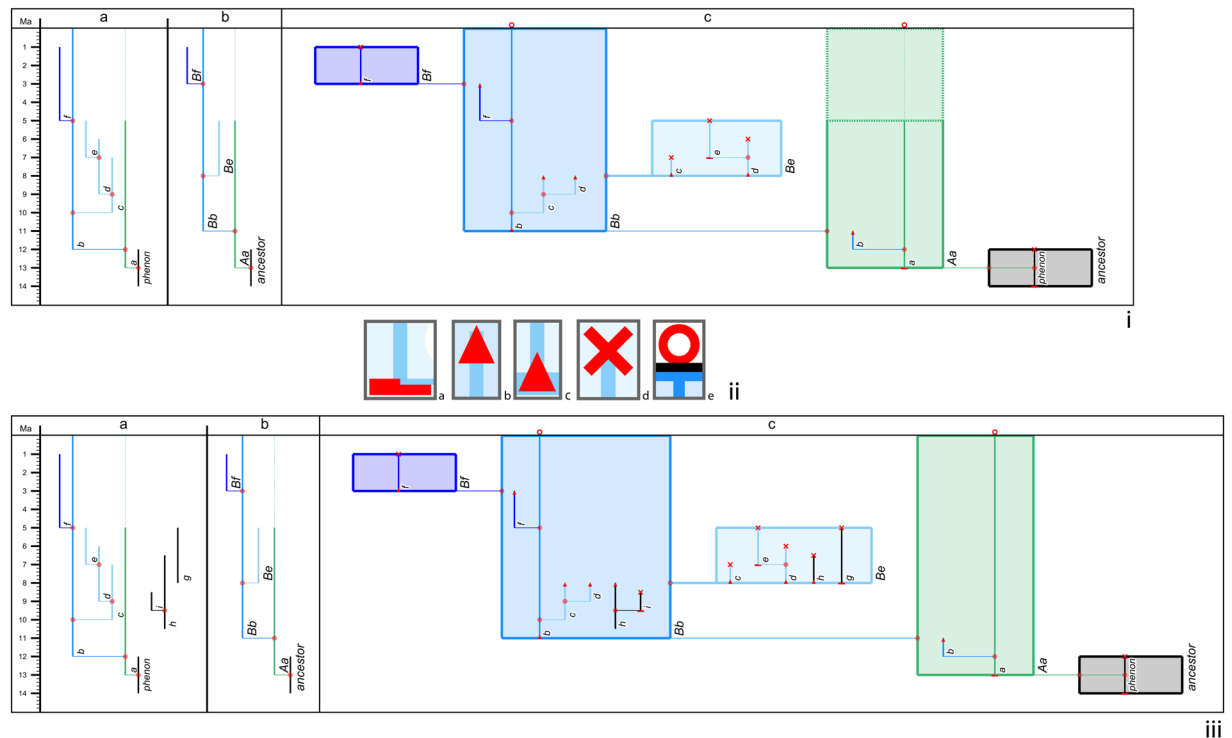


Figure 2. Sample derivation of an integrated species–phenon tree. (i) Improvised example of evolutionary trees against geologic time scale (Ma), depicting a “black” ancestor giving rise to a “green” descendant, in turn to a “blue” group (“medium blue”, ancestral to “light blue” and “dark blue” descendants); all range lines effected in the same line style (“frequent”), except for the upper/later portion of phenon *a* and species *Aa* (“conjectured”; for *TimeScale Creator* line styles, see pp. 48–49 of⁸¹); drawn by *TimeScale Creator* datapacks (§ Data availability). Column a = phenon tree. Column b = species tree. Column c = integrated species–phenon tree. (ii) Enlarged view of guide symbols attached to phenon range lines in column c to provide stratigraphic or phylogenetic context, including their parts broken between species boxes: (a) range origin; (b) top of broken part of range, to transfer to the next descendant species box; (c) bottom of broken part of range, to transfer from the immediate ancestral species box; (d) range extinction; (e) still living. (iii) Incorporation of disconnected phenon trees and ranges (see § Methods); evolutionary-tree series of (i), with added phenon tree (h–i), and additional phenon range (g), all displayed in the same column, from left to right in order of First Occurrence (the default option). Column b: as in (i); the additional black phenon, (g–i), from column a have been assigned to the *Bb*–*Be* lineage series, but their ancestor–descendant relationships with the original coloured phenon are considered poorly known. Column c: integrated species–phenon tree, with the disconnected phenon tree (h–i) and range (g) positioned in their corresponding species boxes.

We now describe the integrated species–phenon tree, apply it to this case example of Cenozoic macroperforate planktonic foraminifera, and then suggest how these trees might enhance broader research into infraspecific diversity, especially that which we seek to detail through deep time.

Results

The software tool. Our integrated species–phenon tree, part of the *TimeScale Creator* software package (§ Methods), is a species tree drawn as a SVG chart against geologic time but with the species expanded into boxes in which is projected an underlying infraspecific (phenon) tree (Fig. 2i). Those infraspecifics could be taxa (sub-species, botanical varieties or forms, prokaryotic infrasubspecifics, etc.) or other entities of interest (molecular OTUs, gene sequences/loci, biostratigraphic taxa, etc.).

The key incentive for a software tool to bring together both species and phenon trees is that, though the trees refer to the same organisms, they have different topologies, in terms of both the nature and number of entities recognized and the timing of their evolution. Even in our simple improvised example of just a few taxa (Fig. 2i), observing which species (column b of Fig. 2i) contain which parts of the phenon tree (column a) is not easy. So our tool does that for you (column c). It also provides range guides (Fig. 2ii) to better appreciate the context of the now broken-up parts of the phenon tree; these make clear the full durations of phenon and easily distinguish these from transfers of the phenon between species. So now the integrated tree can provide not only all the information contained in the individual species and phenon trees, but also enhance understanding of the morphologic/genetic/etc. variability within species, as well as the timing of evolutionary change within and between species, including what morphological change taxa capture within and across speciation events. In so doing it encourages

an examination of the dynamics of morphological evolution in relation to speciation (cladogenesis) and lineage continuity (anagenesis).

Case example. The practical viability of our tool finds a ready check in Aze & others⁵¹ jointly published species and phenon trees for Cenozoic macroperforate planktonic foraminifera, especially apposite as they were recently transferred onto the *TimeScale Creator* platform⁵⁴. As already alluded to above (§ Avoiding “morphospecies”), their trees constitute a major dataset of large and complex trees comprising 210 species (their biological-species lineages) and 339 phenon (their “morphospecies”) over a time interval of 66 Myr. To systematically recognize, let alone name, species lineages is unusual in micropaleontology; Aze & others introduced codes for their lineages, e.g., N133-T135, by concatenating numerical codes given to their included lineage-tree leaves, prefixed with N if internodal, T if terminal. Their “morphospecies”, given as *Genus species* binomina, followed a convention in their field to employ distinctive, especially biostratigraphically useful, morphologies as formal taxa. Moves are underway to revise the trees of Aze & others to incorporate the recently published major changes to Oligocene taxonomy and phylogeny⁷⁶ and analogous but longer-term updates expected from a fledgling Neogene Planktonic Foraminifera Working Group.

Within Aze & others’ trees the rate and content of macroevolutionary change in terms of species and infraspecific taxa is highly variable (Fig. 7 in⁵⁴). The relationship between species and phenon trees, constructed stratophenetically (Fig. 2 in¹), is similarly variable: for example, though origins for about 40% of the species lineages coincide with one of their contained phenon (see “morphospecies” in pdf p. 18 in⁵⁴), many phenon originate within species lineages and a substantial number cross from ancestral to descendant lineages. So envisaging, for example, the phenon content of species lineages from the trees figured in the original paper is just too difficult for the average observer (compare in¹ the top and bottom parts of Fig. 5A–J, or the separately drawn “morphospecies” and lineage trees in Appendices S2 and S3). And this visual challenge is only partially alleviated by enhanced displays (§ Linkages between morphospecies and lineage trees in⁵⁴) introduced to assist with this.

Applying the new integrated species–phenon tree to bring these trees together results in an even wider tree (Fig. 3c); nonetheless the new chart still draws quickly, a testament to the scalability built into the Java code of *TimeScale Creator* software⁷⁷. Closer inspection demonstrates that the phenon range lines (Fig. 3a, inset) have indeed separated at species lineage origins (Fig. 3b, inset) to transfer into their respective species boxes (Fig. 3c, inset). The much easier access to species–phenon relationships afforded by the integrated tree is underlined by comparison of these species boxes with our earlier attempt (Fig. 7f in⁵⁴) at a similar effect by manually overlaying lineage outlines onto the phenon (“morphospecies”) tree. The integrated tree also features range guides (Fig. 3d; legend in Fig. 2ii) to help the viewer appreciate the stratigraphic and phylogenetic context for any phenon range, and especially to follow the trajectory of those that cross species boxes (note the phenon highlighted). In addition, as already stated, the integrated tree brings with it all the information from the individual species and phenon trees. Viewed interactively on the *TimeScale Creator* platform, this includes the mouse-over pop-ups which, in this case example, include a wealth of taxonomic, morphologic, ecologic, and biostratigraphic detail from the back-end database (including the lineage–morphospecies linkages already mentioned)⁵⁴.

Discussion

We have already alluded to the many approaches to visualizing infraspecific content of species through time, so our integrated species–phenon tree has a rich heritage. It could, for instance, be traced back to the then newly enthused evolutionary paleontologists of the late nineteenth century⁷⁸. More specifically, our formulation was anticipated by predigital species–phenon boxes employed to document stratigraphic distributions (Fig. 4b) within the context of a species tree (Fig. 4a). Also, within the context of documenting infraspecific content, there is the “phenon group”, an informal category uniting homologous phenon in multielement fossils^{79,80}. Interestingly, if trees of phenon groups were constructed, given that both descendants at speciation would often retain the same phenon for some phenon groups, the relationship of phenon-group trees with their species trees should bear similarities to relationships between gene trees and their species trees — bringing us back to some topics we raised above.

Coming forward to contemporary research related to our case example of planktonic foraminifera, parallels can be made, especially with regard to living representatives, between the perspective the integrated tree brings and that coming from molecular studies. For instance, the integrated tree explicitly visualizes species as dynamically evolving polytypic lineages, presenting them in terms of a two-level species–phenon hierarchy (Fig. 5a). Meanwhile, some molecular studies are not only genetically circumscribing nominal species and separating out misunderstood cryptic look-a-likes, they are also laying the basis for an informal (but rules-based) multilayered infraspecific taxonomy: a visualization that recognizes the complexity of species lineages and their phylo-geography/ecology/oceanography in terms of genetic lineages, putative species, and populations⁷⁰ (Fig. 5b). Given that deep-time (for example paleoceanographic) applications of planktonic foraminifera depend on developing analogous levels of understanding, one wonders if the species–phenon depiction will continue to be adequate into the future, in regards to either its mere two-level hierarchy or the way phenon (as species-group taxa) are currently employed. However, compared to other phylogenetic schemes currently employed for deep time, the species–phenon tree does provide both a comprehensive and specific set of proposals and timings, in terms of both species and phenon (Fig. 5c). And some of these could potentially be amenable to, for example, molecular dating and phylogenetics. And, despite the rich fossil record of planktonic foraminifera, taking the results of molecular phylogenetics into deep time can usefully suggest parts of this record needing reexamination for stratophenetic phylogenetics (Fig. 5d,e: the case of *Globigerinoides elongata*).

Beyond Aze & others’ phylogeny of planktonic foraminifera, our integrated tree should have immediate application for similar cases where the fossil record is rich enough to motivate the proposal of phylogenetic trees of not just fossil morphotaxa but paleobiologic species. And, if those organisms have living representatives, a similarly fruitful interchange with molecular and other living research would be expected. But we would also encourage

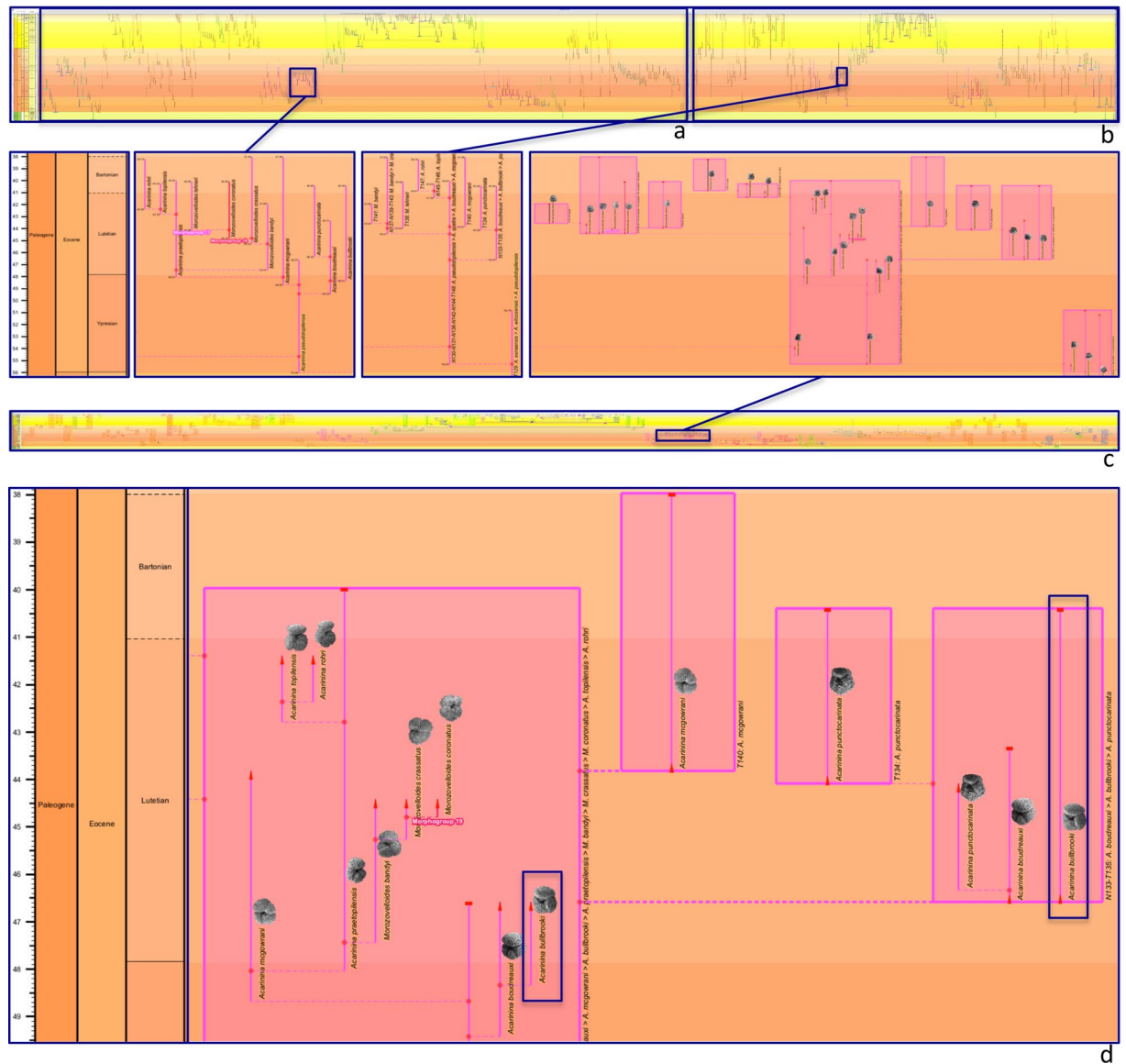


Figure 3. Derivation of an integrated species–phenon tree for the case study of Cenozoic macroperforate planktonic foraminifera¹. Evolutionary-tree charts drawn by *TimeScale Creator* datapacks, (a–c) against entire Cenozoic time scale (the last 66 Myr), with insets of a Ypresian–Bartonian (Eocene) clade which begins with a phenon (morphospecies *Acarinina pseudotopilensis*, a), or with its corresponding species (Lineage N130–N131–N136–N142–N144–T148*, b,c); clade shown within Fig. 5C of¹ and Figs. 7f and 20 of⁵⁴; tree colours and groupings by morphogroup; background coloured by stages (“Chronostrat” option). (a,b) Entire phenon (morphospecies) and species (lineage) trees⁵⁴, respectively, each with inset (below). (c) Entire integrated species–phenon tree, with inset (above); this tree introduced herein (§ Data availability); images⁸⁵ for morphospecies added as heuristics to help appreciate the break-up of the phenon tree within the species tree, but these images are not authoritative. (d) Detail (upper right) from inset of (c); red range guides on phenon range lines as in Fig. 2ii; time interval is late Ypresian–Bartonian (Eocene). Note the breaking graphically of the range line of the highlighted phenon (morphospecies *Acarinina bullbrooki*) between the ancestral species box (Lineage N130–N131–N136–N142–N144–T148, far left) and its descendant species box (Lineage N133–T135, far right). *With regard to species-lineage labels in (c,d), note that, in order to make it easier to follow the lineage codes of Aze & others, as part of the transfer to *TimeScale Creator*⁵⁴ these labels were programmatically appended with a list of included phenon. This label then becomes “N130–N131–N136–N142–N144–T148: *A. pseudotopilensis* > *A. quetra* > *A. boudreauxi* > *A. mcgowrani* > *A. bullbrooki* > *A. praetopilensis* > *M. bandyi* > *M. crassatus* > *A. topilensis* > *A. rohrri*”: this lineage happens to be one of their more inclusive!

researchers with other kinds of temporal infraspecific information to consider placing that within the expandable graphic provided by our tool in order to visualize the phylogenetic context and detail of that information. This information need not be in tree form. For instance, the graphic could be used to display, say, the temporal

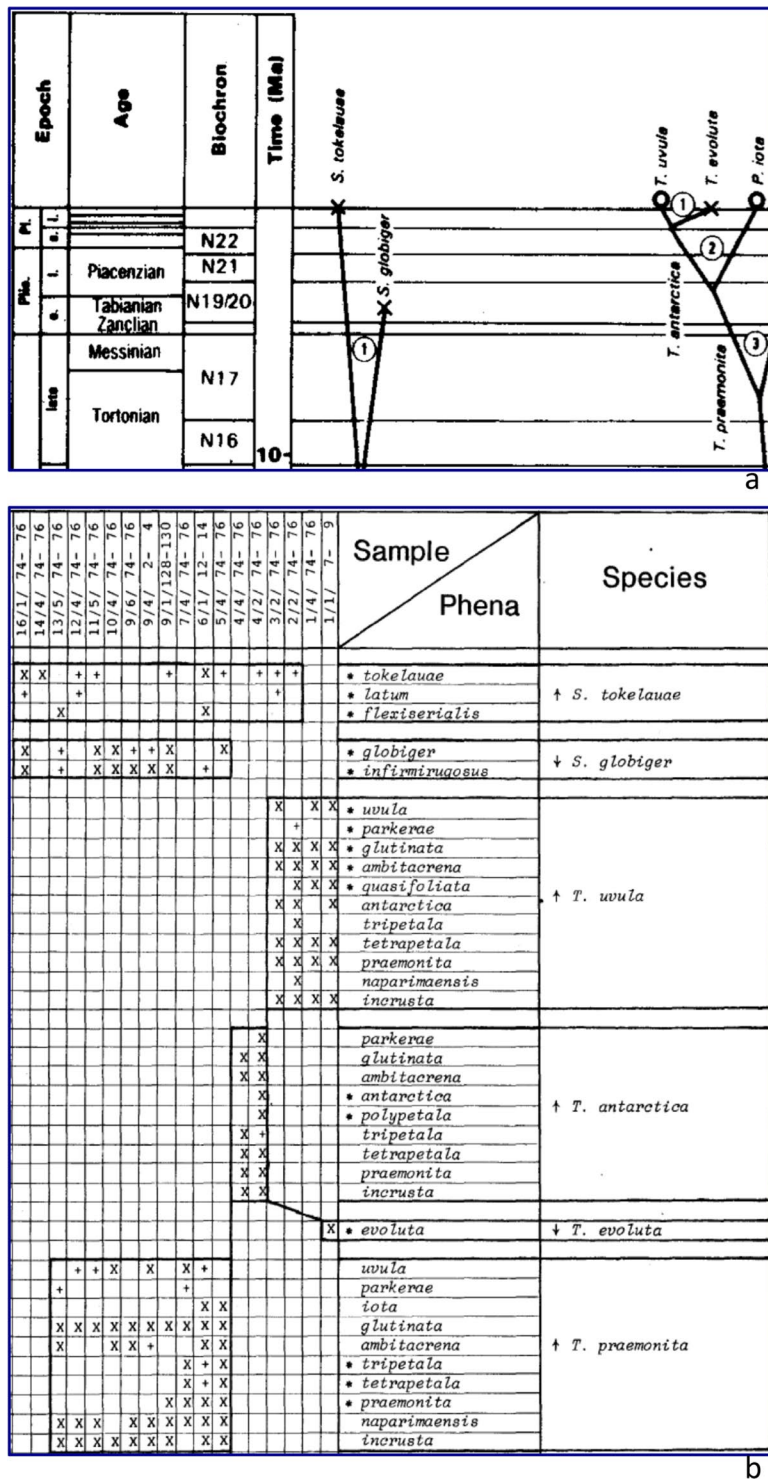


Figure 4. Earlier thinking towards species–phenon trees. A small portion of the phylogeny and stratigraphic charts from a study^{63,86} of Cenozoic planktonic foraminifera: a window limited to the last 10 Myr, focusing in on *Streptochilus* and *Tinophodella praemonita* and descendants (see original for details). (a) Portion of species tree (Text-Fig. 2b of⁶³, upper-left). (b) Corresponding stratigraphic-distribution chart for Deep Sea Drilling Site 208 (Table 3a of⁶³, portion): species boxes arranged and linked according to the species tree, with occurrences of included phena listed by stratigraphic order of primary types.

positions and proposed species memberships of specimens, especially when at the working stage of a phylogenetic study. More generally in a more purely graphical context, a visualization tool which can bring together two nested trees that can nonetheless remain loosely coupled may have broader application.

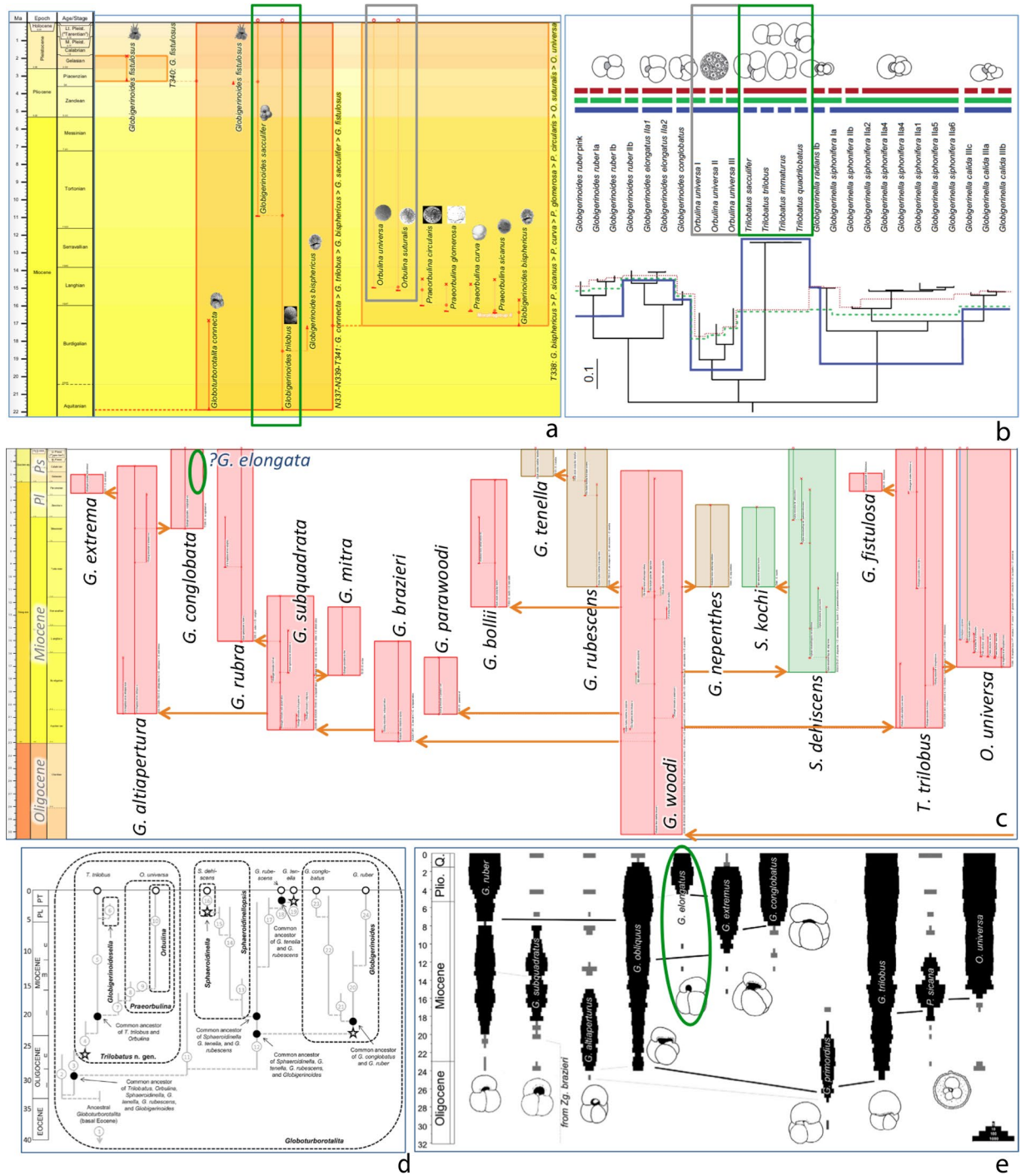


Figure 5. Mutual learnings between integrated species–phenon trees and molecular trees. **(a,b)** Living species of planktonic-foraminifer genera, *Orbulina* and *Trilobatus*, and their evolution during the Neogene. **(a)** Portion with these two genera from the integrated species–phenon tree of the case study of Cenozoic macroperforate planktonic foraminifera (see Fig. 3c). **(b)** A molecular phylogeny of those two genera and related groups (Fig. 3 of⁷⁰). *Orbulina* (grey boxes): rather than a single species lineage (T338) with two phena that are still living, the genetics point to the need for stratophenetic studies to check for progressive speciation through the Miocene–Pleistocene which resulted in three living species. *Trilobatus* (green boxes): stratophenetic studies have already placed the four genetically almost identical nominal species employed by molecular studies into a single species lineage (N337–N339–T341; with two of those nominal species portrayed as intergradational phena, two synonymised). **(c–e)** Phylogenies since the late Oligocene of planktonic foraminifera related to *Globigerinoides*. **(c)** An annotated portion from the integrated species–phenon tree of the case study of Cenozoic macroperforate planktonic foraminifera (see Fig. 3c); coloured by ecogroups; species binomina are not authoritative (genus epithets after⁸⁷); details of phena within species, etc. are viewable interactively when the datapack is loaded onto the *TimeScale Creator* platform (SData availability). **(d)** A stratophenetic scheme influenced by molecular phylogenetics (Fig. 5 of⁸⁷; see for details). **(e)** A molecular–stratigraphic scheme (Fig. 6 of⁴⁹; see for details).

The molecular study (e) suggested species *G. elongata* (green ellipse) is a living descendant of *G. conglobata* and as-yet unrecognized by the stratophenetic approaches [see green ellipse in (c)] as it is a homeomorph of *G. rubra* (though this would be a relatively minor convergence as both share quite a close common ancestor, *G. subquadrata*).

Methods

Introduction and rationale. Our integrated species–phenon tree takes a species tree (Fig. 2i, column b) and phenon tree (column a) for the same organisms, both set against geologic time, and combines the two (column c). This is accomplished by allowing the species time ranges (leaves) of the species tree to expand laterally into rectangles (“boxes”), breaking up the ranges of the phenon tree at the branch points of the species containing them, and then placing the broken-up parts of the phenon tree into their corresponding species boxes. Guide symbols (Fig. 2ii) are added to the extremities of each of the broken phenon ranges so as to indicate their context (their origin, their extinction or still-living presence, or transfer from their ancestral species or to their descendant species).

The new integrated tree is drawn within the Java-based *TimeScale Creator* (*TSCreator*, *TSC*) visualization software package, as were its component species and phenon trees. The integration is achieved with minimal additional data input — a single extra entry for each phenon, denoting the species in which its range ends — the tool does the rest. This minimal indication of links between the two component trees avoids the introduction of errors from manual tree compilation, likely with even the most meticulous effort when the dataset is large (e.g., for our case example¹, the original listings in Appendix S5 of the 339 morphospecies in the 210 lineages contained a small % of such errors, understandably).

Visualization of the integrated tree on the same platform as its component trees makes it easy to transfer all of the latter’s graphical features as well: line styles, colours, thumbnail images, pop-ups, and, most importantly, the exact dimensions of the tree topologies. An additional flexibility built into the integrated-tree tool allows the inclusion of phenon ranges within species boxes for any phenon not connected into the original phenon tree. Similarly, any phenon trees/subtrees disconnected from the main phenon tree will also be drawn within their species boxes; this could be employed for depiction of additional infraspecific categories. A further feature allows the user to provide alternative labels for phenon; this could be employed, for example, to label phenon with only a single species epithet (*species*, of *Genus species*) which is then easily visually distinguished from the binomen of its enveloping species (*Genus species*).

Programming the new integrated tree presented significant challenges. The key issue for the previously coded evolutionary tree for *TimeScale Creator* concerned collation and coordination of widths of diverse graphical components, quite a task in itself. The integrated tree adds further complications arising from the lack of a simple overall parent–child hierarchy, with the result that, for example, the number of breaks within phenon ranges is not easily determined up-front, requiring greater preemptive space complexity and concomitant memory. A key solution has been the computation of a species–phenon relationship map, which is progressively populated by tracking down the phenon tree from top to bottom. With the resulting extended tree data structure, the internal dimensions for species boxes can then be determined and so the graphical footprint of the overarching species tree put in place.

Evolutionary trees in *TimeScale Creator*. *TimeScale Creator* is a Java-based visualization software package which lets users draw interactive charts to explore and compare Earth-history events of any portion of geologic time. Evolutionary trees are now available for these charts (pp. 50–51 of⁸¹). These trees can be displayed in columns alongside other columns such as regional lithostratigraphy, biostratigraphic zonations, isotopic curves, trends, sequences, or global and regional event time-series data. The tool to draw the new integrated species–phenon tree (pp. 52–53 of⁸¹) has been included in this evolutionary-tree feature.

An evolutionary tree is inputted into our software package via a datapack containing a tab-delimited text file. Each row in this file contains text or numeric data within equivalent spreadsheet columns B–K to represent the evolutionary relationship and attributes of a time-range point. The specific format is (pp. 50–51 of⁸¹):

Data row for range

Label [Parent-name] (column B) ⇒ Age (Ma) (C) ⇒ Abundance (D) ⇒ Pop-up(E)

Data row for branch

Parent-name (column B) ⇒ Age (Ma) (C) ⇒ “branch” (D) ⇒ Child-name (E)
 ⇒ “on/off”(F) ⇒ BranchLabel (G) ⇒ Dashed/Dotted (H)
 ⇒ Pop-up (I) ⇒ BranchColor (J) ⇒ Priority (K)

A general tree data structure is a connected network between nodes where a unique path from any ancestor to any descendant is maintained; a reverse path (from descendant to an ancestor) is not allowed. Because we are also positioning evolutionary events against standard geologic time in the evolutionary-tree column, each node in the tree corresponds to an evolutionary range (pp. 48–49 of⁸¹), informing the first-appearance datum (FAD or BASE point of the range) and the last-appearance datum (LAD or TOP point of the range) for an organism. The left–right branching and spacing of descendant ranges starting at the branch point of its parent range is deduced by calculating the number of children (p. 51 of⁸¹) and positioning them on the drawing canvas. Our evolutionary-tree algorithm receives only range data with point and line attributes and then generates the

Relationship ID in the map	Name of species range	Size of phenon range group	Name of grouped phenon ranges	One-to-many relationship map
1-1	ancestor	1	phenon	ancestor ← phenon
2-1	Aa	2	a	Aa ← a
2-2			b	Aa ← b
3-1	Bb	4	b	Bb ← b
3-2			c	Bb ← c
3-3			d	Bb ← d
3-2			f	Bb ← f
4-1	Be	3	c	Be ← c
4-2			d	Be ← d
4-3			e	Be ← e
5-1	Bf	1	f	Bf ← f

Table 1. Species–phenon relationship map for the integrated tree of Fig. 2i.

tree data structure determining the branch connections among parent and child ranges and hierarchies among subtrees.

The main challenge in drawing an evolutionary tree is the determination of the left–right horizontal and up–down vertical positioning of the range lines in order to avoid collisions between branches and subtrees and so generate a visually balanced tree. Our algorithm determines the width of left and right subtrees branching out from the parent ranges to recursively calculate the total width of the final tree. It draws the vertical range lines on the drawing canvas first, followed by the horizontal branch lines.

Integrated species–phenon tree: additional programming. Drawing of the integrated tree adds a further layer of programming to that of the evolutionary tree. In particular, the width determination and positioning algorithm now needs to accommodate the breaking up of phenon range lines into parts which are then transferred across multiple rectangular species range boxes. For example, in Fig. 2i we can see that the phenon range *b* is broken into two range-line segments, the first in the species range box *Aa* and the second in *Bb*. Similarly, parts of the phenon ranges *c* and *d* are transferred from the species range box *Bb* to *Be*. This latter instance more obviously demonstrates that this breaking-up and transfer process disrupts the phenon tree structure and gives rise to disconnected subtrees within descendant species range boxes in the integrated tree. There are two new subtrees inside the species box *Be* in Fig. 2i (column c), the first one contains only the partial range line *c* and the second one is a subtree rooted at the range line *d* inside the species box *Be*. The integrated-tree algorithm needs to address the additional spacing and information transfer for these new partial ranges forming these segregated subtrees.

Programming for the integrated tree also needs to handle the possibility of phenon ranges split between several species boxes forming a successive series of ancestors and descendants. A single phenon range may be transferred to more than two species boxes depending on its first and last appearance point. The transferring of the same contents of a phenon range to its broken-up parts across multiple species boxes adds to the complications. It requires more sophisticated data structure, higher space complexity, and greater memory consumption, because there is not a fixed limit on the number of breaks in phenon range lines which may be needed. The simple counting of children per parent range under each subtree and subsequent width determination in the existing tree-drawing algorithm are not sufficient. Besides the widths of range lines and subtrees, the widths of the rectangular species range boxes need to be precisely calculated after the extended set of range lines are created. For this reason, the algorithm needs to map species–phenon relationships between an individual species range and its occupant group of phenon ranges (Table 1).

Integrated species–phenon tree: additional data to link phenon and species. In order to link and combine the species and phenon trees, additional information needs to be provided in the datapack. However, our program minimizes this additional input by exploiting the convenient but somewhat surprising graphical fact that only one linking item is needed to accomplish this: for each phenon we need only provide the species in which its range ends. Once the top of a phenon time range is positioned in its corresponding species box, the earlier parts of its tree can be progressively transferred onto the species tree algorithmically (see previous subsection). This insight is especially handy for the user as it minimizes not only the effort needed to gather the additional data required but also the potential for errors to creep in if additional lists of mutual species–phenon memberships are compiled by other means (see § Introduction and rationale, above).

To generate an extended tree data structure the format of the input dataset of the datapack needs to be modified. The TOP range point of each phenon range line is employed to distinguish between species and phenon range lines; the tree algorithm uses this to separately create species and phenon trees. Only two entries are added to the TOP range point of an individual phenon range: (a) column G now contains the flag “phenon” and (b) column H contains the name of the species range box in which this TOP range point of the current phenon range will be positioned (see also dataset TSCEvolTree_IntTree2019Fig. 2ic.txt, § Data availability):

Data row for TOP range point

Label (column B) ⇒ Age (Ma) (C) ⇒ TOP (D) ⇒ Pop-up(E) ⇒ blank (F)
 ⇒ “phenon” (G) ⇒ Species-name (H)

Integrated species–phenon tree: programming implementation. Given two side-by-side trees, the first one a species tree and the second one a phenon tree, we can establish the species–phenon mapping relationship by simply knowing which species range box will include the last appearance datum (LAD or TOP) of a phenon range. The new algorithm starts from the LAD or TOP point of a phenon range, associates this TOP point with its occupying species range line using the “phenon” flag (column H in the new format) and maps the first relationship, then continues comparing subsequent earlier (geologically past) range point ages of the phenon range line with the points of the mapped species range line. Every time it finds a younger branch connection age in the species tree than the phenon range (FAD point of the current species range greater than the phenon range points), it ends the current phenon range line under consideration by adding a new first appearance datum (FAD or BASE point). It simultaneously transfers the rest of the phenon range line to the parent species box of the current box after adding a new last appearance datum (LAD or TOP point) to the phenon range. Similarly, it continually adds new TOP and BASE points for the broken-up phenon range lines and populates the species–phenon relationship maps (Table 1) until it reaches the BASE point of the starting phenon range. This same process is applied to all the phenon range lines and, in the end, the algorithm acquires the new extended set of range lines with new range points for the phenon range lines and branch connections for further resolution of width, new symbols, and colour inheritance.

Each species range from the species tree expands horizontally to a box in the integrated tree (Fig. 2i). Each box has the exact time range (i.e., vertical extent) as on the species tree, and the branches adopted from the species tree are extended horizontally. Species boxes are basically the same species ranges with a rectangular shape and varied widths. Therefore, the boxed species tree maintains the same topology as the species tree before integration and phenon-range transfer. The species box is filled with colour that matches the colour of the inherited branch label, but with some added transparency (in the example from Fig. 2i, species box *Aa* is green but with a RGBA alpha value chosen to enable the phenon range lines, labels, and images to be visible and readable). When a colour is assigned to a branch in the datapack dataset, the ranges of its descendants inherit that colour until a new colour is assigned to a later descendant.

The choice of line styles for the range (see Fig. 2) of the species is exactly transferred to the box. In the example in Fig. 2i, all ranges except one are “frequent”, so all boxes except one have all sides with a “frequent” thickness. The exception is *Aa*, which is “frequent” from 13 to 5 Ma and “conjectured” from 5 Ma to the present. So the box is drawn with a “frequent” thickness along its base and 8 Myr up its sides from the base, and with a “conjectured” top and then 5 Myr down the sides from the top. Ranges can be assigned any of several abundance settings for any part of a range, so the equivalent species box will need to reflect this.

The width of each species box is dependent on the number of phenon ranges inside it. Using the species–phenon relationship map, the algorithm quickly resolves the group of phenon ranges to be drawn inside each species box. It takes consideration of the width of both the range label and image for each phenon range. It is also possible that there may exist a species box without phenon ranges; these are given a minimum width to handle such exceptions.

Each instance of the phenon range retains its core features. When phenon splits are required, all the features of the phenon ranges are duplicated and attached to each broken-up portion, e.g., the thumbnail image and pop-ups. However, the line style of the phenon ranges strongly follows the parent phenon tree because line styles are directly associated with the abundance details (line styles) for the phenon range. Visual distinction between species and phena is accomplished by employing different settings for the “default” line styles (frequent vs. common, respectively). The species range box boundaries are therefore thicker than the phenon range line (Fig. 2i, column c).

In order to make it easier to appreciate the context of phenon range lines, especially to distinguish between their ranges and their parts broken between species boxes, guide symbols are added to the extremities of each of their parts (Fig. 2ii).

Integrated species–phenon tree: disconnected trees and phenon ranges. Our integrated-tree tool can manage multiple phenon trees (Fig. 2iii: original coloured tree, with additional phenon tree *h–i* in black). If the datapack contains more than one mutually unconnected tree, the preexisting evolutionary-tree algorithm has been retained to draw all trees within the same *TimeScale Creator* column, positioning them side by side based on “First occurrence”, “Last occurrence” or “Alphabetic order” (p. 49 of⁸¹).

Our integrated tree can also handle disconnected phenon ranges: not all of the phenon ranges need to be part of a tree structure (Fig. 2iii: additional phenon *g* in black). The algorithm orders the phenon ranges using the “First Occurrence” by default. But three options, (a) First Occurrence, (b) Last Occurrence, and (c) Alphabetic Order, are provided to the user from the settings window of the software. Ordering of the phenon ranges inside the species box can thus be freed from the topology of the unified phenon tree.

Integrated species–phenon tree: short phenon labels. The integrated tree allows alternative labels for phena, usually to conserve space. Another application for this could be to apply a particular nomenclatural format for phenon labels. For instance, if a study conferred scientific names on phena, one could opt for phena to be given only a single species epithet (*species*, of *Genus species*), and so distinguish them from binomina (*Genus species*) of their enveloping species boxes. This feature is implemented by placing “|” within the phenon range

name in the dataset, in which case the text after “[” is excluded; if the name does not contain “[”, then the full name is used. (This option was effected in the back-end database of the case study via the table SpeciesGroupName, in which the original assignments of scientific names are rearranged to begin with the most trivial name — in order to provide an index invariant to later nomenclatural revisions; see links to details of the database in Table 2 and Fig. 8 of⁵⁴. For example, to take morphospecies *Acarinina bullbrooki*, featured in Fig. 3d, its original assignment, “*Globorotalia bullbrooki* Bolli, 1957”, rearranges to “*bullbrooki* Bolli, 1957; *Globorotalia*”. If then the phenon range name in the dataset is given as “*bullbrooki* | Bolli, 1957; *Globorotalia*”, the phenon range label displayed on the integrated tree will be reduced to “*bullbrooki*”). If the user’s key concern is to more distinctly distinguish visually between species and phenon labels, the fonts for each can be separately manipulated (*TimeScale Creator* settings: Choose Time Interval, select datapack column, Fonts).

Data availability

The *TimeScale Creator* datapacks introduced herein for the integrated species–phenon tree from the case study of Cenozoic macroperforate planktonic foraminifera are available at Australian National University Data Commons collection anudc:5981, <https://doi.org/10.25911/5db66faba683b82>.

Datapacks

Coloured by ecogroup

With labels: TSCevolTree_Aze&2011_CorrJul2018_ISPEco.dpk

Without labels: TSCevolTree_Aze&2011_CorrJul2018_ISPEcoNoLbl.dpk

Coloured by morphogroup

With labels: TSCevolTree_Aze&2011_CorrJul2018_ISPMph.dpk

Without labels: TSCevolTree_Aze&2011_CorrJul2018_ISPMphNoLbl.dpk

Settings: TSCevolTree_Aze&2011_CorrJul2018_ISP_4dpks.tsc

Note that these datapacks do not include the images of foraminifera featured in Fig. 3, because authoritative images are not yet available for these datasets. These datapacks are also freely available from the *TimeScale Creator* website <https://timescalecreator.org/datapack/datapack.php>

Other Supplementary Information for this article is available at Australian National University Data Commons collection anudc:5982, <https://doi.org/10.25911/5db66fd8c512783>.

TimeScale Creator datapacks for Fig. 2i

Datasets

TSCevolTree_IntTree2019Fig. 2ia,b.txt

TSCevolTree_IntTree2019Fig. 2ic.txt

Datapacks

TSCevolTree_IntTree2019Fig. 2ia,b.dpk

TSCevolTree_IntTree2019Fig. 2ic.dpk

Settings

TSCevolTree_IntTree2019Fig. 2ia,b.tsc

TSCevolTree_IntTree2019Fig. 2ic.tsc

TimeScale Creator datapacks for Fig. 2iii

Datasets

TSCevolTree_IntTree2019Fig. 2iiia,b.txt

TSCevolTree_IntTree2019Fig. 2iiic.txt

Datapacks

TSCevolTree_IntTree2019Fig. 2iiia,b.dpk

TSCevolTree_IntTree2019Fig. 2iiic.dpk

Settings

TSCevolTree_IntTree2019Fig. 2iiia,b.tsc

TSCevolTree_IntTree2019Fig. 2iiic.tsc

“Functional Specification for the Evolutionary Trees function, providing an option for an Integrated Species–Phenon Tree, 27 August 2014”: Functional Specification _ Integrated Species – Phenon Tree.docx.

TimeScale Creator datapacks, the relational database, and related files for the previously published lineage and morphospecies trees from the case study are also available at Australian National University Data Commons collections (for details see⁵⁴). These datapacks are also freely available from the *TimeScale Creator* website.

Code availability

The source-code files for the Integrated Species–Phenon Tree feature are available at https://github.com/brishtiteveja/TSCreator_Integrated_Evolutionary_Tree and at

Australian National University Data Commons collection anudc:5980, <https://doi.org/10.25911/5db66f7bc193584>. These repositories contain the source-code files constituting the core programming for the Integrated Species–Phenon Tree feature within the Evolutionary Tree column of the *TimeScale Creator* software platform.

RangeColumn.java: includes detailed data structures for evolutionary range points and ranges, and tree-drawing utility functions.

ImageGenerator.java: draws the evolutionary tree data column and generates the final chart.

Main.java: includes the main function which is the entry point to the software and spawns the Java process on a Java Virtual Machine.

Readme.txt.

The *TimeScale Creator* software platform is freely available at <https://timescalecreator.org/download/download.php>, from where it can be downloaded as a Java archive (jar) file or Windows executable (exe) file. Instructions to install *TimeScale Creator* and load datapacks are given in *Instructions_for_TSCreator_datapacks.pdf*, included in anudc:5980.

Received: 7 June 2019; Accepted: 25 October 2019;

Published online: 12 December 2019

References

- Aze, T. *et al.* A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data. *Biological Reviews of the Cambridge Philosophical Society* **86**, 900–927, <https://doi.org/10.1111/j.1469-185X.2011.00178.x> (2011).
- Etienne, R. S. *et al.* Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proceedings of the Royal Society. B, Biological sciences* **279**, 1300–1309, <https://doi.org/10.1098/rspb.2011.1439> (2012).
- Darwin, C. *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life.* (John Murray, 1859).
- Bouzat, J. L. Darwin's diagram of divergence of taxa as a causal model for the origin of species. *The Quarterly Review of Biology* **89**, 21–38, <https://doi.org/10.1086/674992> (2014).
- de Queiroz, K. Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences of the USA* **102**(Suppl 1), 6600–6607, <https://doi.org/10.1073/pnas.0502030102> (2005).
- Gulla, S., Lund, V., Kristoffersen, A. B., Sørum, H. & Colquhoun, D. J. *vapA* (A-layer) typing differentiates *Aeromonas salmonicida* subspecies and identifies a number of previously undescribed subtypes. *Journal of Fish Diseases* **39**, 329–342, <https://doi.org/10.1111/jfd.12367> (2016).
- Hardion, L. *et al.* Does infraspecific taxonomy match species evolutionary history? A phylogeographic study of *Arundo formosana* (Poaceae). *Botanical Journal of the Linnean Society* **183**, 236–249, <https://doi.org/10.1093/botlinnean/bow006> (2017).
- Bernini, G. *et al.* Complexity of biogeographic pattern in the endangered crayfish *Austropotamobius italicus* in northern Italy: molecular insights of conservation concern. *Conservation Genetics* **17**, 141–154, <https://doi.org/10.1007/s10592-015-0767-4> (2016).
- Serrano-Villavicencio, J. E., Vendramel, R. L. & Siniciato Terra Garbino, G. Species, subspecies, or color morphs? Reconsidering the taxonomy of *Callicebus* Thomas, 1903 in the Purus–Madeira interfluvium. *Primates* **58**, 159–167, <https://doi.org/10.1007/s10329-016-0555-x> (2017).
- Fujimoto, T. *et al.* Biochemical, physiological, and molecular characterization of *Dickeya dianthicola* (formerly named *Erwinia chrysanthemi*) causing potato blackleg disease in Japan. *Journal of General Plant Pathology* **84**, 124–136, <https://doi.org/10.1007/s10327-018-0772-9> (2018).
- Belfiore, N. M., Liu, L. & Moritz, C. Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: Geomyidae). *Systematic Biology* **57**, 294–310, <https://doi.org/10.1080/10635150802044011> (2008).
- Weir, B. S., Johnston, P. R. & Damm, U. The *Colletotrichum gloeosporioides* species complex. *Studies in mycology* **73**, 115–180, <https://doi.org/10.3114/sim0011> (2012).
- Barrett, R. A. *et al.* A chloroplast phylogeny of *Zieria* (Rutaceae) in Australia and New Caledonia shows widespread incongruence with species-level taxonomy. *Australian Systematic Botany* **27**, 427–449, <https://doi.org/10.1071/SB14033> (2015).
- Wen, D., Yu, Y., Hahn, M. W. & Nakhleh, L. Reticulate evolutionary history and extensive introgression in mosquito species revealed by phylogenetic network analysis. *Molecular Ecology* **25**, 2361–2372, <https://doi.org/10.1111/mec.13544> (2016).
- Sota, T., Kojima, T., Lee, Y. J. & Lin, C.-P. Phylogenetic relationships of Japanese *Auritibicen* species (Hemiptera: Cicadidae: Cyrotympanini) inferred from mitochondrial and nuclear gene sequences. *Zoological Science* **33**, 401–406 (2016).
- Han, T., Lee, W., Lee, S., Park, I. G. & Park, H. Reassessment of species diversity of the Subfamily Denticollinae (Coleoptera: Elateridae) through DNA barcoding. *PLoS One* **11**, e0148602, <https://doi.org/10.1371/journal.pone.0148602> (2016).
- Bogarín, D. *et al.* Anchored hybrid enrichment generated nuclear, plastid and mitochondrial markers resolve the *Lepanthes horrida* (Orchidaceae: Pleurothallidinae) species complex. *Molecular Phylogenetics and Evolution* **129**, 27–47, <https://doi.org/10.1016/j.ympev.2018.07.014> (2018).
- Mizuyama, M., Masucci, G. D. & Reimer, J. D. Speciation among sympatric lineages in the genus *Palythoa* (Cnidaria: Anthozoa: Zoantharia) revealed by morphological comparison, phylogenetic analyses and investigation of spawning period. *PeerJ* **6**, e5132, <https://doi.org/10.7717/peerj.5132> (2018).
- Cuezzo, M. G., Miranda, M. J., Vogler, R. E. & Beltramino, A. A. From morphology to molecules: a combined source approach to untangle the taxonomy of *Clessinia* (Gastropoda, Odontostomidae), endemic land snails from the Dry Chaco ecoregion. *PeerJ* **6**, e5986, <https://doi.org/10.7717/peerj.5986> (2018).
- Luo, A., Zhu, C.-D., Ho, S. Y. W. & Ling, C. Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology* **67**, 830–846, <https://doi.org/10.1093/sysbio/syy011> (2018).
- Gaubert, P., Taylor, P. J., Fernandes, C. A., Bruford, M. W. & Veron, G. Patterns of cryptic hybridization revealed using an integrative approach: a case study on genets (Carnivora, Viverridae, *Genetta* spp.) from the southern African subregion. *Biological Journal of the Linnean Society* **86**, 11–33, <https://doi.org/10.1111/j.1095-8312.2005.00518.x> (2005).
- Charruau, P. *et al.* Phylogeography, genetic structure and population divergence time of cheetahs in Africa and Asia: evidence for long-term geographic isolates. *Molecular Ecology* **20**, 706–724, <https://doi.org/10.1111/j.1365-294X.2010.04986.x> (2011).
- Lancaster, M. L., Arnould, J. P. Y. & Kirkwood, R. Genetic status of an endemic marine mammal, the Australian fur seal, following historical harvesting. *Animal Conservation* **13**, 247–255, <https://doi.org/10.1111/j.1469-1795.2009.00325.x> (2010).
- Emami-Khoyi, A. *et al.* Mitochondrial DNA structure and colony expansion dynamics of New Zealand fur seals (*Arctocephalus forsteri*) around Banks Peninsula. *New Zealand Journal of Zoology* **43**, 322–335, <https://doi.org/10.1080/03014223.2016.1179649> (2016).
- Allende, C., Sohn, E. & Little, C. Treelink: data integration, clustering and visualization of phylogenetic trees. *BMC Bioinformatics* **16**, e414, <https://doi.org/10.1186/s12859-015-0860-1> (2015).
- Benavente, E. D. *et al.* PhyTB: phylogenetic tree visualisation and sample positioning for *M. tuberculosis*. *BMC Bioinformatics* **16**, e155, <https://doi.org/10.1186/s12859-015-0603-3> (2015).
- Nei, M. *Molecular evolutionary genetics.* (Columbia University Press, 1987).
- Doyle, J. J. Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany* **17**, 144–163, <https://doi.org/10.2307/2419070> (1992).
- Maddison, W. P. Gene trees in species trees. *Systematic Biology* **46**, 523–536, <https://doi.org/10.2307/2413694> (1997).
- Avise, J. C. & Robinson, T. J. Hemiplasy: a new term in the lexicon of phylogenetics. *Systematic Biology* **57**, 503–507, <https://doi.org/10.1080/10635150802164587> (2008).
- Goodman, M., Czelusniak, J., Moore, G. W., Romero-Herrera, A. E. & Matsuda, G. Fitting the gene lineage into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences. *Systematic Zoology* **28**, 132–163, <https://doi.org/10.2307/2412519> (1979).
- Szöllösi, G. J., Tannier, E., Daubin, V. & Boussau, B. The inference of gene trees with species trees. *Systematic Biology* **64**, e42–e62, <https://doi.org/10.1093/sysbio/syu048> (2014).

33. Kim, N. & Lee, C. Three-Dimensional Phylogeny Explorer: distinguishing paralogs, lateral transfer, and violation of “molecular clock” assumption with 3D visualization. *BMC Bioinformatics* **8**, e213, <https://doi.org/10.1186/1471-2105-8-213> (2007).
34. Scornavacca, C., Huson, D. H. & Zickmann, F. Tanglegrams for rooted phylogenetic trees and networks. *Bioinformatics* **27**, i248–i256, <https://doi.org/10.1093/bioinformatics/btr210> (2011).
35. Singh, S. P., Groeneveld, J. C., Al-Marzouqi, A. & Willows-Munro, S. A molecular phylogeny of the spiny lobster *Panulirus homarus* highlights a separately evolving lineage from the southwest Indian Ocean. *PeerJ* **5**, e3356, <https://doi.org/10.7717/peerj.3356> (2017).
36. Hofmeyr, M. D., Vamberger, M., Branch, W., Schleicher, A. & Daniels, S. R. Tortoise (Reptilia, Testudinidae) radiations in southern Africa from the Eocene to the present. *Zoologica Scripta* **46**, 389–400, <https://doi.org/10.1111/zsc.12223> (2017).
37. Hills, S. F. K., Crampton, J. S., Trewick, S. A. & Morgan-Richards, M. DNA and morphology unite two species and 10 million year old fossils. *PLoS One* **7**, e52083, <https://doi.org/10.1371/journal.pone.0052083> (2012).
38. Vitek, N. S. Delineating modern variation from extinct morphology in the fossil record using shells of the Eastern Box Turtle (*Terrapene carolina*). *PLoS One* **13**, e0193437, <https://doi.org/10.1371/journal.pone.0193437> (2018).
39. Armstrong, H. A. & Brasier, M. D. *Microfossils*. (Blackwell, 2005).
40. Knappertsbusch, M. Morphological variability of *Globorotalia menardii* (planktonic foraminifera) in two DSDP cores from the Caribbean Sea and the eastern equatorial Pacific. *Carnets de Géologie, Article 7* http://paleopolis.rediris.es/cg/CG2007_A04/index.html (2007).
41. Schiebel, R. & Hemleben, C. *Planktic foraminifers in the modern ocean*. (Springer, 2017).
42. Huber, B. T., Bijma, J. & Darling, K. Cryptic speciation in the living planktonic foraminifer *Globigerinella siphonifera* (d’Orbigny). *Paleobiology* **23**, 33–62, <https://doi.org/10.1017/S0094837300016638> (1997).
43. de Vargas, C. & Pawlowski, J. Molecular versus taxonomic rates of evolution in planktonic foraminifera. *Molecular Phylogenetics and Evolution* **9**, 463–469, <https://doi.org/10.1006/mpev.1998.0491> (1998).
44. Darling, K. F., Wade, C. M., Kroon, D., Brown, A. J. L. & Bijma, J. The diversity and distribution of modern planktic foraminiferal small subunit ribosomal RNA genotypes and their potential as tracers of present and past ocean circulations. *Paleoceanography* **14**, 3–12, <https://doi.org/10.1029/1998PA900002> (1999).
45. de Vargas, C., Renaud, S., Hilbrecht, H. & Pawlowski, J. Pleistocene adaptive radiation in *Globorotalia truncatulinoides*: genetic, morphologic, and environmental evidence. *Paleobiology* **27**, 104–125, [http://dx.doi.org/10.1666/0094-8373\(2001\)027<0104:PARIGT>2.0.CO;2](http://dx.doi.org/10.1666/0094-8373(2001)027<0104:PARIGT>2.0.CO;2) (2001).
46. de Vargas, C., Bonzon, M., Rees, N. W., Pawlowski, J. & Zaninetti, L. A molecular approach to biodiversity and biogeography in the planktonic foraminifer *Globigerinella siphonifera* (d’Orbigny). *Marine Micropaleontology* **45**, 101–116, [https://doi.org/10.1016/S0377-8398\(02\)00037-3](https://doi.org/10.1016/S0377-8398(02)00037-3) (2002).
47. Darling, K. F., Kucera, M., Pudsey, C. J. & Wade, C. M. Molecular evidence links cryptic diversification in polar planktonic protists to Quaternary climate dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 7657–7662, <https://doi.org/10.1073/pnas.0402401101> (2004).
48. Sexton, P. F. & Norris, R. D. Dispersal and biogeography of marine plankton: long-distance dispersal of the foraminifer *Truncorotalia truncatulinoides*. *Geology* **36**, 899–902, <https://doi.org/10.1130/G25232A.1> (2008).
49. Aurahs, R., Treis, Y., Darling, K. & Kucera, M. A revised taxonomic and phylogenetic concept for the planktonic foraminifer species *Globigerinoides ruber* based on molecular and morphometric evidence. *Marine Micropaleontology* **79**, 1–14, <https://doi.org/10.1016/j.marmicro.2010.12.001> (2011).
50. Weiner, A. K. M. *et al.* Phylogeography of the tropical planktonic foraminifera lineage *Globigerinella* reveals isolation inconsistent with passive dispersal by ocean currents. *PLoS One* **9**, e92148, <https://doi.org/10.1371/journal.pone.0092148> (2014).
51. Weiner, A. K. M., Weinkauff, M. F. G., Kurasawa, A., Darling, K. F. & Kucera, M. Genetic and morphometric evidence for parallel evolution of the *Globigerinella calida* morphotype. *Marine Micropaleontology* **114**, 19–35, <https://doi.org/10.1016/j.marmicro.2014.10.003> (2015).
52. Ujiie, Y. & Ishitani, Y. Evolution of a planktonic foraminifer during environmental changes in the tropical oceans. *PLoS One* **11**, e0148847, <https://doi.org/10.1371/journal.pone.0148847> (2016).
53. Pearson, P. N. A lineage phylogeny for the Paleogene planktonic foraminifera. *Micropaleontology* **39**, 193–232, <https://doi.org/10.2307/1485898> (1993).
54. Fordham, B. G. *et al.* Future-proofing the Cenozoic macroperforate planktonic foraminifera phylogeny of Aze & others (2011). *PLoS One* **13**, e0204625, <https://doi.org/10.1371/journal.pone.0204625> (2018).
55. de Vargas, C., Zaninetti, L., Hilbrecht, H. & Pawlowski, J. Phylogeny and rates of molecular evolution of planktonic foraminifera: SSU rDNA sequences compared to the fossil record. *Journal of Molecular Evolution* **45**, 285–294, <https://doi.org/10.1007/pl00006232> (1997).
56. Darling, K. F. & Wade, C. M. The genetic diversity of planktic foraminifera and the global distribution of ribosomal RNA genotypes. *Marine Micropaleontology* **67**, 216–238, <https://doi.org/10.1016/j.marmicro.2008.01.009> (2008).
57. Ujiie, Y. & Lipps, J. H. Cryptic diversity in planktic foraminifera in the northwest Pacific Ocean. *Journal of Foraminiferal Research* **39**, 145–154, <https://doi.org/10.2113/gsjfr.39.3.145> (2009).
58. Ujiie, Y., de Garidel-Thoron, T., Watanabe, S., Wiebe, P. & de Vargas, C. Coiling dimorphism within a genetic type of the planktonic foraminifer *Globorotalia truncatulinoides*. *Marine Micropaleontology* **77**, 145–153, <https://doi.org/10.1016/j.marmicro.2010.09.001> (2010).
59. Quillévéré, F. *et al.* Global scale same-specimen morpho-genetic analysis of *Truncorotalia truncatulinoides*: a perspective on the morphological species concept in planktonic foraminifera. *Palaeogeography, Palaeoclimatology, Palaeoecology* **391**, 2–12, <https://doi.org/10.1016/j.palaeo.2011.03.013> (2013).
60. Blow, W. H. Late middle Eocene to Recent planktonic foraminiferal biostratigraphy. In *Proceedings of the First International Conference on Planktonic Microfossils* (eds Bronnimann, P. & Renz, H. H.) Vol. 1, 199–421 (E. J. Brill, 1969).
61. Sprovieri, A., Ruggieri, G. & Untl, M. *Globorotalia truncatulinoides excelsa* n. subsp., foraminifero planctonico guida per il Pleistocene inferiore. *Bollettino della Società Geologica Italiana* **99**, 3–11 (1980).
62. Bolli, H. M. & Saunders, J. B. Oligocene to Holocene low latitude planktic foraminifera. In *Plankton stratigraphy* (eds Bolli, H. M., Saunders, J. B. & Perch-Nielsen, K.) Vol. 1, 155–262 (Cambridge University Press, 1985).
63. Fordham, B. G. *Miocene–Pleistocene planktic foraminifers from D. S. D. P. Sites 208 and 77, and phylogeny and classification of Cenozoic species*. 200 pp., 25 pls. (Evolutionary Monographs, 6, <https://www.mn.uio.no/cees/english/services/van-valen/evolutionary-monographs/>, 1986).
64. Stewart, D. R. M. *Evolution of Neogene globorotaliid foraminifera and Miocene climate change*. (PhD thesis, University of Bristol, <https://research-information.bristol.ac.uk/files/34489299/288306.pdf>) (2003).
65. Saito, T., Breger, D. & Thompson, P. R. *Systematic index of Recent and Pleistocene planktonic foraminifera*. (University of Tokyo Press, 1981).
66. Kennett, J. P. & Srinivasan, M. S. *Neogene planktonic foraminifera: a phylogenetic atlas*. (Hutchinson Ross, 1983).
67. Cifelli, R. & Scott, G. H. Stratigraphic record of the Neogene globorotalid radiation (planktonic Foraminiferida), <https://doi.org/10.5479/si.00810266.58.1>. *Smithsonian Contributions to Paleobiology* **58** (1986).
68. Lazarus, D., Hilbrecht, H., Spencer-Cervato, C. & Thierstein, H. Sympatric speciation and phyletic change in *Globorotalia truncatulinoides*. *Paleobiology* **21**, 28–51, <https://doi.org/10.1017/S0094837300013063> (1995).

69. Chaisson, W. P. & Pearson, P. N. Planktonic foraminifer biostratigraphy at Site 925: middle Miocene–Pleistocene. In *Proceedings of the Ocean Drilling Program. Scientific Results* (eds Shackleton, N. J., Curry, W. B., Richter, C. & Bralower, T. J.) 154, 3–32, http://www-odp.tamu.edu/publications/154_SR/01_CHAP.PDF (1997).
70. Morard, R. *et al.* Nomenclature for the nameless: a proposal for an integrative molecular taxonomy of cryptic diversity exemplified by planktonic foraminifera. *Systematic Biology* **65**, 925–940, <https://doi.org/10.1093/sysbio/syw031> (2016).
71. Mayr, E. *Principles of systematic zoology*. (McGraw-Hill, 1969).
72. Mayr, E. *The growth of biological thought*. (Belknap Press, 1982).
73. Smith, A. B. *Systematics and the fossil record*. (Blackwell Scientific, 1994).
74. Wurst, E. & Frank, B. Contributions to a revision of the genus *Schwiebea* (Acari: Acaridae). I. Redescription of *Schwiebea talpa* and *Schwiebea nesbitti*. *Stuttgarter Beiträge Zur Naturkunde, Serie A* **579** (1998).
75. Wieland, B. *et al.* Phenon cluster analysis as a method to investigate epidemiological relatedness between sources of *Campylobacter jejuni*. *Journal of Applied Microbiology* **100**, 316–324, <https://doi.org/10.1111/j.1365-2672.2005.02788.x> (2006).
76. Wade, B. S., Olsson, R. K., Pearson, P. N., Huber, B. T. & Berggren, W. A. (Eds.) *Atlas of Oligocene planktonic foraminifera*. (Cushman Foundation for Foraminiferal Research Special Publication, vol. **46**, 2018).
77. Ogg, J. G. & Gradstein, F. M. TS-Creator—chronostratigraphic data base and visualisation. Cenozoic–Mesozoic–Paleozoic integrated stratigraphy and user-generated time scale graphics and charts. *Episodes* **29**, 65–66 <https://doi.org/10.18814/epiugs/2006/v29i1/008> (2006).
78. Archibald, J. D. *Aristotle's ladder, Darwin's tree: the evolution of visual metaphors for biological order*. (Columbia University Press, 2014).
79. Fordham, B. G. A literature-based phylogeny and classification of Silurian conodonts. *Palaeontographica, Abteilung A* **217**, 1–136 (1991).
80. Fordham, B. G. Advantages of Comprehensive Taxonomy for routine systematic documentation in conodont biostratigraphy. *Boletín de la Academia Nacional de Ciencias, Córdoba [Argentina]* **60**, 469–482 (1995).
81. Ogg, J. G. & Ogg, G. M. (Eds.) *TimeScale Creator Manual*. (Geologic TimeScale Foundation Inc., https://timescalecreator.org/download/TS-Creator%20Manual_July2019.zip) (2019).
82. Zehady, A. K., Fordham, B. G. & Ogg, J. G. *Cenozoic macroperforate planktonic foraminifera phylogeny of Aze & others (2011). TimeScale Creator Integrated Species–Phenon Tree. Corrected Version, July 2018. Four datapacks for Java software package*. (Australian National University Data Commons Collection anudc:5981, <https://doi.org/10.25911/5db66faba683b>) (2019).
83. Zehady, A. K., Fordham, B. G. & Ogg, J. G. *Integrated species–phenon trees: visualizing infraspecific diversity within lineages. Data repository for published article*. (Australian National University Data Commons Collection anudc:5982, <https://doi.org/10.25911/5db66fd8c5127>) (2019).
84. Zehady, A. K. *Source code, Integrated Species–Phenon Tree feature, Evolutionary Tree column, TimeScale Creator visualization platform*. (Australian National University Data Commons Collection anudc:5980, <https://doi.org/10.25911/5db66f7bc1935>) (2019).
85. Haller, C. & Ogg, J. G. Cenozoic macroperforate planktonic foraminifera evolution [Poster]. *Third Geological Problem Solving with Microfossils Conference (Microfossils III), 10–13 March, 2013, North American Micropaleontology Section (NAMS) of the Society for Sedimentary Geology (SEPM)* (2013).
86. Fordham, B. G. *A chronostratigraphy and corresponding classification of Neogene planktonic foraminifera with a documentation of morphotypes from two Pacific deep-sea stratigraphic sections*. (PhD thesis, University of Queensland, <https://openresearch-repository.anu.edu.au/handle/1885/9007>) (1979).
87. Spezzaferri, S. *et al.* Fossil and genetic evidence for the polyphyletic nature of the planktonic foraminifera “*Globigerinoides*”, and description of the new genus *Trilobatus*. *PLoS One* **10**, e0128108, <https://doi.org/10.1371/journal.pone.0128108> (2015).

Acknowledgements

Tracy Aze (University of Leeds), Paul Pearson (Cardiff University), and Bridget Wade (University College London) generously helped us appreciate our case example in other studies. Major improvements to the manuscript resulted from two thoughtful and astute anonymous reviews, respectfully coordinated by Editorial Board Member, Mark Chapman. Zixiao Wang and Genevieve Turner (Australian National University: ANU) guided our data onto the ANU Data Commons. Stephen M. Eggins (ANU) provided valuable intellectual input and institutional resources. Funding for the software development was part of an educational grant to A.K.Z. from the non-profit Geologic TimeScale Foundation. Publication of this article was fully funded by Purdue University Libraries Open Access Publishing Fund, coordinated by Nina Collins (Purdue University Press).

Author contributions

A.K.Z. designed, developed, and tested the graphics program, and wrote the methods sections of the paper. B.G.F. conceived the visualization goals, wrote the functional specifications, augmented the database and program to generate the datapacks, wrote the main text of the paper, and arranged for software and data curation. J.G.O. oversaw the design of the visualization package, supervised the development of the programming, and arranged online datapack access. B.G.F. and A.K.Z. prepared the figures. A.K.Z., J.G.O., and B.G.F. reviewed and edited the paper.

Competing interests

Each of the authors declared that there is not any potential competing interest, financial or non-financial.

Additional information

Correspondence and requests for materials should be addressed to B.G.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019